DOCUMENT RESUME

ED 190 667 TM 800 471

AUTHOR Allen, Russell: Fredrick, Vicki
TITLE Testing Services Manual.
INSTITUTION Wisconsin State Dept. of Public Instruction, Madison. Div. for Management and Planning Services.
REPORT NO WSDPI-Bull-0241
PUB DATE [79]
NOTE 37p.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Standards: Computer Assisted Testing: Consumer Economics: *Criterion Referenced Tests: *Educational Assessment: Educational Objectives: *Educational Testing: Elementary Secondary Education: *Item Banks: Language Arts: Mathematics: Reading Skills: Scoring: Test Construction: *Testing Programs
IDENTIFIERS Rasch Model: Wisconsin: *Wisconsin Pupil Assessment Program

ABSTRACT

The following steps have been taken to accomplish the goals of the Wisconsin Pupil Assessment Program: the development of objective-referenced tests: the administration of a standardized norm-referenced test: reporting of trends in performance on these tests: and the availability of two testing services and personnel for consultation purposes. These two services are fully described in this manual. One of the services allows schools to administer the same objective-referenced tests as those used by the state assessment program, and the other allows schools to construct tests using an item bank. Advantages, development, content areas, objectives, and scoring reports of both the statewide and customized tests are defined and described. Also discussed is general testing information, including such areas as sampling/testing of all students, setting performance standards, utilizing and disseminating test data, and considerations in reporting to three target groups: teachers, school board members, and students and their parents. (Author/GK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

TESTING SERVICES MANUAL

prepared by
Russell Allen and Vicki Fredrick

DIVISION FOR MANAGEMENT,
PLANNING AND FEDERAL SERVICES

A. A. Buchmiller, Assistant Superintendent

Bureau of Educational Planning, Evaluation,
Information and Research
Tom Stefonek, Director

Wisconsin Pupil Assessment Program
Russell Allen
Diedre Brensike
Vicki Fredrick
Darwin Kaufman
Susan Ketchum
Angie Wisinger

Bulletin No. 0241

APR 4 1980

TABLE OF CONTENTS

INTRODUCTION 1
STATEWIDE TESTS 2
Advantages 2
Type of Tests 2
Development 2
Content Areas 2
Mathematics 3
Reading 3
Writing 3
Economic Understanding 3
Table 1: Statewide Tests 4
Objectives 4
Test Objectives in Mathematics 5
Fourth Grade Mathematics Test:
Computation and Applications 5
Computation 5
Money Applications 5
Measurement Applications 5
Fourth Grade Mathematics Test:
Geometry and Number Sentences 5
Geometry 5
Mathematical Sentences 6
Eighth and Twelfth Grade Mathematics
Tests: Computation and Applications
Computation 6
Money Applications 6
Measurement Applications 6
Eighth and Twelfth Grade Mathematics
Test: Geometry 7
Size and Shape 7
Points and Sets of Points 7
Symmetry 7
Congruence 7
Similarity 7
Coordinate Systems and Graphs 7
Test Objectives in Reading 8
Fourth Grade Reading Test: Educational
Reading 8
Fourth Grade Reading Test: Everyday Reading 8
Eighth Grade Reading Test: Educational
Reading 8
Eighth Grade Reading Test: Everyday Reading . . . 8
Twelfth Grade Reading Test: Educational
Reading 9
Twelfth Grade Reading Test: Everyday Reading . . . 9
Test Objectives in Writing 10
Fourth Grade Writing Test: Social Writing 10
Fourth Grade Writing Test: Scholastic Writing . . 10

Eighth Grade Writing Test: Social Writing 10
Eighth Grade Writing Test: Business/Vocational Writing 10
Eighth Grade Writing Test: Scholastic Writing . . . 10
Eleventh Grade Writing Test: Social Writing 10
Eleventh Grade Writing Test: Business/Vocational Writing 11
Eleventh Grade Writing Test: Scholastic Writing . . 11
Test Objectives in Economic Understanding 11
Economic Understanding—Consumer Economics 11
Economic Understanding—Traditional Economics . . . 11
Scoring Reports for Mathematics, Reading, and Economic Understanding 12
Scoring Reports for Writing 12
CUSTOMIZED TESTS 14
Advantages 14
Figure 1: Average RIT Scores for Reading and Mathematics . . . 15
Type of Tests 16
Development 16
Content Areas 17
Objectives 17
Using the Item Bank 17
Step 1: Develop Objectives 17
Step 2: Match Objectives to Course Goals 17
Step 3: Select Items 17
Step 4: Review Items 17
Step 5: Format Test Booklets 17
Step 6: Print Copies of Tests 17
Scoring Reports 18
GENERAL TESTING INFORMATION 19
Sampling or Testing All Students 19
Setting Performance Standards 19
What Choices Are There in Setting Standards? 19
Table 2: Determining Sample Size from a Given Population Size 20
What is to Be Gained from Setting Performance Standards Prior to Obtaining Test Results? 21
What Sorts of General Considerations Need to Be Made in Setting Performance Standards? 21
Selecting a Committee 22
Choosing the Objectives Which Are to Have Standards 22
Method I 23
Method II 25
Two Standard Setting Methods Which Are Not Reportable by The Statewide Test Computer Scoring System 27
Method III 27
Method IV 29
A Concluding Remark Concerning Performance Standards 30
Utilization and Dissemination of the Test Data 30
Some General Considerations 30
Three Possible Target Groups 31
Teachers 31
School Board Members 32
Students and Their Parents 32

Introduction

The Wisconsin Pupil Asssessment Program has four major goals:

1. To provide the legislature, governor's office, and general citizenry of Wisconsin with a profile of how well a sample of Wisconsin public school students demonstrate expected skills and knowledge.

   To accomplish this first goal, the Pupil Assessment Program has developed several objective-referenced tests.

2. To provide the legislature, the governor's office, and general citizenry of Wisconsin with a profile of how well a sample of Wisconsin public school students perform in comparison to the rest of the nation.

   The Pupil Assessment Program administers a standardized norm-referenced test to satisfy this goal.

3. To provide the legislature, the governor's office, and general citizenry of Wisconsin with periodic progress reports that indicate systematic upward or downward trends in performance.

   Systematic trends in performance on the objective-referenced tests and the norm-referenced test are reported by the Pupil Assessment Program.

4. To provide materials and technical assistance to schools which voluntarily utilize the products and processes of the Pupil Assessment Program.

   In order to accomplish this goal, the Pupil Assessment Program offers two testing services and personnel for consultation purposes. One service allows schools to administer the same objective-referenced tests as those used by the state assessment program. The other service allows schools to construct their own tests using items contained in an item bank.

The two testing services available to schools from the Pupil Assessment Program are described in this manual. Because costs vary slightly from year to year, it is not possible to list exact prices in this manual. For the most recent cost information or for consultation with staff, contact the Pupil Assessment office at (608)267-7268 or (608)266-3126.

Statewide Tests

Schools may use the existing objective-referenced tests which have been developed for statewide assessment, instructions for administering the tests, test scoring arrangements, and reports of test results provided by the Wisconsin Pupil Assessment Program. This testing service offers several advantages to schools.

Advantages

- Large expenditures of teacher time for test development are eliminated. The tests are readily available for use.

- The tests, having been developed specifically for testing in Wisconsin, may be more relevant to the needs of Wisconsin schools than tests developed for national distribution.

- Schools may set performance standards for their students, basing the standards on the curriculum and expectations of the individual schools or district.

- State averages are available for all tests. Schools may compare their students' scores with the state averages if they test the same grade levels.

Type of Tests

All statewide tests offered within this service are objective-referenced. Objective-referenced tests are designed to measure performance in terms of specific objectives or skills. The statewide tests are multiple-choice tests, with the exception of the writing tests, which require written samples.

Development

Each of the tests available from the Wisconsin Pupil Assessment Program has gone through a lengthy developmental process. Wisconsin educators and non-educators played a major role in identifying objectives, selecting and writing items, and reviewing the items after pilot testing. Tests were also reviewed and revised after each assessment.

Content Areas

Tests have been developed in four areas: Mathematics, Reading, Writing, and Economic Understanding. All of the statewide objective-referenced tests are designed to assess both academic and practical skills.

Although these tests were developed for the specific grade levels mentioned in this manual, they may be used at other grade levels. Schools should review items and objectives to determine at what grade level the tests can be most effectively used.

Mathematics

Grade 4

Two fourth grade tests have been developed. One test assesses Computation and Application skills; the other assesses Geometry and Number Sentence skills.

Grades 8 and 12

One test, assessing Computation and Application skills, has been developed for use at both eighth and twelfth grades. A second test, assessing geometry skills, has been developed for use at both eighth and twelfth grades.

Reading

Grades 4, 8, and 12

A separate test has been developed at each of the three grade levels. Each test consists of two parts: Educational Reading and Everyday Reading. All of the reading tests assess reading comprehension skills.

Writing

All of the writing tests require student writing samples. It is recommended that the testing period for the writing tests be no longer than 30 minutes.

Grade 4

The fourth grade test consists of two parts: Social Writing and Scholastic Writing. Each part contains two items, both requiring student writing samples.

Grades 8 and 11

A separate test has been developed at each of the two grade levels. Each test consists of three parts: Social Writing, Business/Vocational Writing, and Scholastic Writing. Each part contains two items.

Economic Understanding

Grade 12

At grade twelve a test of economic understanding is available. The test consists of two parts: Consumer Economics and Traditional Economics.

Table 1: Statewide Tests

The table below lists all test forms, showing the number of objectives, number of test items, and estimated administration time.

Grade Level*	Test Form	No. of Objectives	No. of Items	Estimated Time Needed to Complete Test Form**
4	Computation/Application	18	72	45-50 minutes
4	Geometry and Number Sentences	13	49	45-50 minutes
4	Educational Reading	4	35	45-50 minutes
4	Everyday Reading	9	36	45-50 minutes
4	Social Writing Skills	2	2	15-20 minutes
4	Scholastic Writing Skills	2	2	25-30 minutes
8	Computation/Application	23	90	45-50 minutes
8	Geometry	16	62	45-50 minutes
8	Educational Reading	9	36	45-50 minutes
8	Everyday Reading	11	44	45-50 minutes
8	Social Writing Skills	2	2	15-20 minutes
8	Business/Vocational Writing Skills	2	2	10-15 minutes
8	Scholastic Writing Skills	2	2	25-30 minutes
11	Social Writing Skills	2	2	15-20 minutes
11	Business/Vocational Writing Skills	2	2	10-15 minutes
11	Scholastic Writing Skills	2	2	25-30 minutes
12	Computation/Application	23	90	45-50 minutes
12	Geometry	16	62	45-50 minutes
12	Educational Reading	15	60	45-50 minutes
12	Everyday Reading	9	36	45-50 minutes
12	Consumer Economics	9	36	20-30 minutes
12	Traditional Economics	12	48	45-50 minutes

*Although all tests are identified with specific grade levels, they can be administered to students at other grade levels. Schools should carefully review test forms to determine suitability.

**Although students should be given as much time as needed to complete a test, nearly all students will finish any of the multiple-choice tests within one class period (45-50 minutes) and any of the writing tests within 30 minutes.

Objectives

The following is a list of objectives assessed by the 1979 statewide tests. Although the objectives are reviewed and revised each year, they tend to remain essentially the same.

TEST OBJECTIVES IN MATHEMATICS

Fourth Grade Mathematics Test: Computation and Applications

Computation

1. Add whole numbers without regrouping.
2. Add whole numbers with regrouping.
3. Subtract whole numbers without regrouping.
4. Subtract whole numbers with regrouping.
5. Multiply whole numbers with basic facts.
6. Multiply whole numbers beyond basic facts, no regrouping.
7. Multiply whole numbers beyond basic facts, with regrouping.
8. Divide whole numbers using ÷ symbol with basic facts.
9. Divide whole numbers using ⟌ symbol with basic facts.
10. Divide whole numbers using the division algorithm, without regrouping or remainder.

Money Applications

11. Solve simple, pictorial problems to find total costs and total coins.
12. Solve multi-step pictorial problems to find total cost, tax, change and earnings.
13. Solve simple problems presented without a picture to find total cost and change from a purchsse.
14. Solve multi-step problems presented without a picture to find total cost, earnings and change.

Measurement Applications

15. Solve simple pictorial problems to find total weight, perimeter; read graph or ruler.
16. Solve multi-step pictorial problems to find total miles, area and differences in temperature.
17. Solve simple problems presented without pictures to find total time, length and capacity.
18. Solve multi-step problems presented without pictures to find total time, units and weight.

Fourth Grade Mathematics Test: Geometry and Number Sentences

Geometry

1. Size and shape: Identify pictures and terms of two-dimensional figures; classify and apply concepts of size and shape.
2. Sets and points: Identify pictures, classify and note relationships of sets and points.
3. Symmetry: Recognize, classify and apply concepts of symmetry.
4. Congruence: Identify congruent figures and apply to real life situations.

5. Similarity: Identify similar figures, recognize the relationship between similar and congruent figures and apply the concept to real world objects.
6. Coordinate systems: Number points on lines of coordinate planes, and apply concept to real life situations.

Mathematical Sentences

7. Language: Use the terms less than, greater than and equal to.
8. Properties: Replace missing numbers in open sentences.
9. Properties: Place +, -, x, or ÷ in placeholders of sentences.
10. Solve number sentences which use the special properties of whole numbers.
11. Solve number sentences using deduction.
12. Identify mathematical statements which describe realistic situations.
13. Solve simple number sentences which represent physical situations.

Eighth and Twelfth Grade Mathematics Test: Computation and Applications

Computation

1. Add and subtract whole numbers.
2. Multiply whole numbers.
3. Divide whole numbers.
4. Find percents of whole numbers.
5. Find the percent one whole number is of another.
6. Find percent-fraction-decimal equivalent.
7. Add terminating decimals.
8. Subtract terminating decimals with regrouping.
9. Multiply decimals with regrouping.
10. Divide decimals with terminating quotients.
11. Use ratio and proportion to solve equivalent fractions.
12. Add and subtract fractions with like denominators.
13. Add and subtract fractions with unlike denominators.
14. Multiply fractions using a horizontal form.
15. Divide using fractions.

Money Applications

16. Solve problems related to purchases (cost, possible purchase, change, discount, tax or surcharge).
17. Solve problems dealing with earnings and savings.
18. Solve problems dealing with basic living and/or traveling expenses.

Measurement Applications

19. Solve problems dealing with length or area.
20. Solve problems dealing with weight and/or volume.
21. Solve simple and complex problems dealing with rate, time and/or temperature.
22. Read graphs and maps.
23. Read and interpret tables.

Eighth and Twelfth Grade Mathematics Test: Geometry

Size and Shape

1. Identify drawings of solids and polygons.
2. Classify angles, triangles, quadrilaterals and other polygons.
3. Deduce a one-step conclusion when given a circle, triangle, quadrilateral, solid and/or other information.

Points and Sets of Points

4. Identify a point, line segment, polygon, radius or diameter; Classify lines as parallel, perpendicular or intersecting.
5. Relationships: Identify supplementary, complementary, vertical and adjacent angles; and relationships between points, lines and planes.
6. Use known properties to deduce conclusions about plane and solid figures; and apply concepts about sets of points to realistic situations.

Symmetry

7. Identify figures with one line of symmetry.
8. Identify figures with more than one line of symmetry, image points; and symmetrical real life objects.

Congruence

9. Identify congruent angles, polygons and corresponding parts of plane figures.
10. Use properties to deduce conclusions about congruent triangles and polygons.

Similarity

11. Identify similar triangles, quadrilaterals, and plane figures.
12. Use properties to deduce conclusions about similar figures.
13. Apply notions of similarity to real life situations.

Coordinate Systems and Graphs

14. Use ordered pairs of whole numbers to identify a point, line, triangle or quadrilateral.
15. Use coordinates to identify the relationship between a point and a given figure.
16. Identify the coordinates and/or the points which fit a given description.

TEST OBJECTIVES IN READING

Fourth Grade Reading Test: Educational Reading

From examples of materials typically read by fourth graders, the student will be able to . . .

1. answer questions about time sequence, events, places and characters in a passage.
2. answer questions about characterizations in a narrative selection.
3. answer questions about the cause of events in a passage.
4. identify the main idea in a passage.

Fourth Grade Reading: Everyday Reading

5. recognize appropriate information from a set of rules.
6. recognize appropriate information from a set of directions.
7. recognize appropriate information from a health warning.
8. recognize appropriate information from a telephone directory page.
9. recognize appropriate information from a dictionary page.
10. recognize appropriate information from a table of contents or index from a textbook.
11. recognize appropriate information from common street and highway signs.
12. recognize appropriate information from a highway map.
13. recognize appropriate information from a graph or chart.

Eighth Grade Reading Test: Educational Reading

From examples of materials typically read by eighth graders, the student will be able to . . .

1. answer questions about time sequence, events, places and characters in a narrative selection.
2. answer questions about the cause of events or actions in a newspaper or magazine selection.
3. identify the main idea in a newspaper or magazine selection.
4. identify the main idea in an expository selection.
5. identify a generalization statement that is appropriate for an expository selection.
6. identify a writer's point of view in a newspaper editorial or magazine selection.
7. separate fact from opinion in a newspaper or magazine article.
8. separate fact from opinion in an advertisement.
9. identify the implied cause and effect of an event described in a newspaper or magazine selection.

Eighth Grade Reading Test: Everyday Reading

10. recognize appropriate information from a set of rules.
11. recognize appropriate information from medical instructions.
12. recognize appropriate information from a health warning.

13. recognize appropriate information from a container label with instructions, warnings or comparative buying information.
14. recognize appropriate information from a telephone directory page.
15. recognize appropriate information from a dictionary page.
16. recognize appropriate information from a product warranty.
17. recognize appropriate information from a recipe.
18. recognize appropriate information from a business letter.
19. recognize appropriate information from a graph or chart.
20. recognize appropriate information from want ads or job descriptions.

Twelfth Grade Reading Test: Educational Reading

From examples of materials typically read by twelfth graders, the student will be able to . . .

1. answer questions about time sequence, events, places and characters in a narrative selection.
2. answer questions about time sequence, events, places and characters in an expository selection.
3. answer questions about time sequence, events, places and characters in a poem.
4. answer questions about the cause of events or actions in a newspaper or magazine selection.
5. answer questions about the cause of events or actions in an expository selection.
6. answer questions about the cause of events or actions in a narrative selection.
7. identify the main idea in a newspaper or magazine selection.
8. identify the main idea in a poem.
9. identify the generalization statement that is appropriate for a newspaper or magazine selection.
10. identify the generalization statement that is appropriate for an expository selection.
11. identify the generalization statement that is appropriate for a narrative selection.
12. answer questions about implied facts in a newspaper or magazine selection.
13. answer questions about implied times, events, or actions in a narrative selection.
14. separate fact from opinion when reading a newspaper or magazine article such as an editorial.
15. identify bias in a newspaper or magazine selection.

Twelfth Grade Reading Test: Everyday Reading

16. recognize appropriate information from operating instructions for an appliance.
17. recognize appropriate information from a business letter.
18. recognize appropriate information from a container label with instructions, warnings, or medical antidotes.
19. recognize appropriate information from a library catalog card.
20. recognize appropriate information from a dictionary page.

21. recognize appropriate information from a voting ballot.
22. recognize appropriate information from a computerized bill.
23. recognize appropriate information from want ads or job descriptions.
24. recognize appropriate comparative buying information from warranties, ratings, advertising, or container labels.

TEST OBJECTIVES IN WRITING

Fourth Grade Writing Test: Social Writing

The student will be able to . . .

1. write a friendly letter.
2. give clear, accurate and complete written directions on how to do or make something.

Fourth Grade Writing Test: Scholastic Writing

The student will be able to . . .

3. write a descriptive essay.
4. write an expository essay.

Eighth Grade Writing Test: Social Writing

The student will be able to . . .

1. give clear, accurate and complete written directions on how to do or make something.
2. write a message that will be clear to the receiver.

Eighth Grade Writing Test: Business/Vocational Writing

3. write a letter requesting a product presented in an advertisement.
4. fill out an application blank neatly and correctly.

Eighth Grade Writing Test: Scholastic

The student will be able to . . .

5. write a descriptive essay.
6. write an expository essay.

Eleventh Grade Writing Test: Social Writing

The student will be able to . . .

1. give clear, accurate and complete written direction on how to do or make something.
2. write a message that will be clear to the receiver.

Eleventh Grade Writing Test: Business/Vocational Writing

The student will be able to . . .

3. write a letter of application for a job presented in a newspaper advertisement.
4. fill out an application blank neatly and correctly.

Eleventh Grade Writing Test: Scholastic Writing

5. write a descriptive essay.
6. write a persuasive essay.

TEST OBJECTIVES IN ECONOMIC UNDERSTANDING

Economic Understanding: Consumer Economics

1. Students should understand how the consumer can affect what is produced in the American economic system.
2. Students should understand the concept of insurance in terms of financial planning and risk management (e.g., health, social, life, liability and property insurance).
3. Students should understand the concept of credit, including types and costs.
4. Students should understand the benefits and risks of various types of savings and investments.
5. Students shall evaluate alternative choices when given a set of circumstances relating to needs and want (e.g., food, clothing, shelter, transportation, recreation, etc.).
6. Students shall identify and be able to use a variety of sources of information about products and services (e.g., catalogs, consumer guides, institutions, yellow pages, newspapers, and personal contracts).
7. Students shall interpret and evaluate advertising (e.g., amount of real savings, appeal to pressure tactics, representation of products and savings, etc.).
8. Students shall be able to use information in selecting goods and services (e.g., packaging, labeling, guarantees, warranties, etc.).
9. Students shall identify a financial plan for spending and saving that is comparable with one's personal values, goals and resources.

Economic Understanding: Traditional Economics

10. Students should understand that because we have limited resources and the potential for unlimited wants there is a need to make choices and to economize.
11. Students should understand the concepts of supply and demand.
12. Students should understand the concept of collective bargaining.
13. Students should understand the concept of inflation.
14. Students should understand the American economic system and how it compares with other economic systems.

15. Students should understand the effect that current consumption patterns may have on the environment and future availability and utilization of resources--both for the individual and for the larger society.
16. Students should understand the causes and effects of international trade.
17. Students should understand the role of tax (fiscal) and monetary (money) policies in relationship to employment and price stability.
18. Students should understant that income, including profits and wages, is generated through the production and exchange of goods and services.
19. Students should understand that the government modifies the market economy by intervention and regulation.
20. Students should be able to interpret economic data, as found in graphs, tables, rations, index numbers, etc.
21. Students should understand that specialization increases efficiency leads to interdependence.

Scoring Reports for Mathematics, Reading, and Economic Understanding

Schools using the Mathematics, Reading, and Economic Understanding tests developed by the Wisconsin Pupil Assessment Program can receive three types of scoring reports:

1. District-School Profile -- lists all objectives and the items measuring each objective; indicates, for each item, the percent of students who select the correct response.

2. Individual Pupil Profile -- indicates how well individual students perform on each objective and each item; if performance standards are set, shows whether or not the student meets the standards (see "Setting Performance Standards" on page 19).

3. Item Tally Profile by School and District -- shows percent of students selecting each possible response for each item.

Scoring Reports for Writing

The writing tests ask students to produce actual writing samples. These must be hand-scored using one of the following two methods:

1. Primary-trait Scoring
   Primary-trait scoring is a method of hand-scoring essays to determine if the way in which they are written successfully accomplishes the main purpose of a specified writing task.

   The term is based on the assumption that particular writing tasks require particular approaches and no others if they are to be successful. The approach used by the writer to reach and affect his/her audiences will be the most important (the primary) trait of a piece of writing. For example, the writer of a set of directions must present things in a logical and unambiguous manner if he/she expects readers to follow the directions. The primary trait of a written set

of directions would be an unambiguous, sequential and logical progression of instructions. Successful papers will have that trait; unsuccessful papers will not, regardless of how clever or well-written they may be in other respects. Each task requiring primary-trait scoring will be rated by two readers per essay. The readers will rate the papers on a scale of 1-4, according to predetermined criteria for each task. A "0" will indicate no response or response not related to the task, and an "X" will indicate an illegible or illiterate response. A third reader will resolve any discrepancies in scores.

2. Holistic Scoring

Holistic scoring is a method of hand-scoring essays to determine their relative quality. The term is derived from this method's emphasis on a reader's response to the whole essay--how well the student communicates. Readers will be trained by rating papers exemplifying a range of score points (scale of 1-4), until they rate papers consistently. They then read the essays, giving each a score within the predetermined scale, based on their overall reaction to a paper. A third reader will resolve any discrepancies.

The following objectives will be scored holistically. All other objectives will be scored by the primary-trait method.

4th Grade

-The student will be able to write a descriptive essay.
-The student will be able to write an expository essay.

8th Grade

-The student will be able to write a descriptive essay.
-The student will be able to write an expository essay.

11th Grade

-The student will be able to write a descriptive essay.

Scoring reports will consist of school averages by grade and a listing of individual student scores on each writing sample.

Customized Tests

Schools may construct tests using items contained in an item bank. Scoring of the tests and score reporting are also part of this service.

It is important to recognize that this item bank represents more than a mere collection of items. Each item is related to specific course objectives and has been assigned a difficulty level, using a statistical procedure developed by Gorg Rasch*, a Danish mathematician. The difficulty levels, referred to as Rasch Units (abbreviated as RITs), are three-digit numbers ranging from about 150 RITs to about 250 RITs (covering the approximate ability range of students in grades 3-8). As shown in Figure 1, fifth graders average about 200 RITs on the reading items and about 200 RITs on the mathematics items; eighth graders average about 217 RITs on the reading items and about 221 RITs on the mathematics items.

This customized testing service offers several advantages to schools.

Advantages

-Each school can design tests which contain objectives and items matched to its individual curriculum.

-The difficulty level of a test can be designed to suit the ability level of a student so that he/she is not frustrated by items that are too difficult or not challenged by items that are too easy.

For example, suppose one wishes to test a class of fifth grade students in reading. Three separate tests could be created--for students of low, average, and high ability. The "low" ability test might consist of items with difficulty levels between 180-200; the "average" ability, 190-200; and the "high" ability, 200-220. In effect, the testing has been "individualized", yet the test scores can be interpreted as though each student took the same test. This is because the Rasch scoring procedure has been shown to yield approximately the same scores for students, regardless of the average difficulty level of the tests taken by the students. This is the case unless a test is far too difficult for the student, so that he or she is required to guess the correct answer to most items.

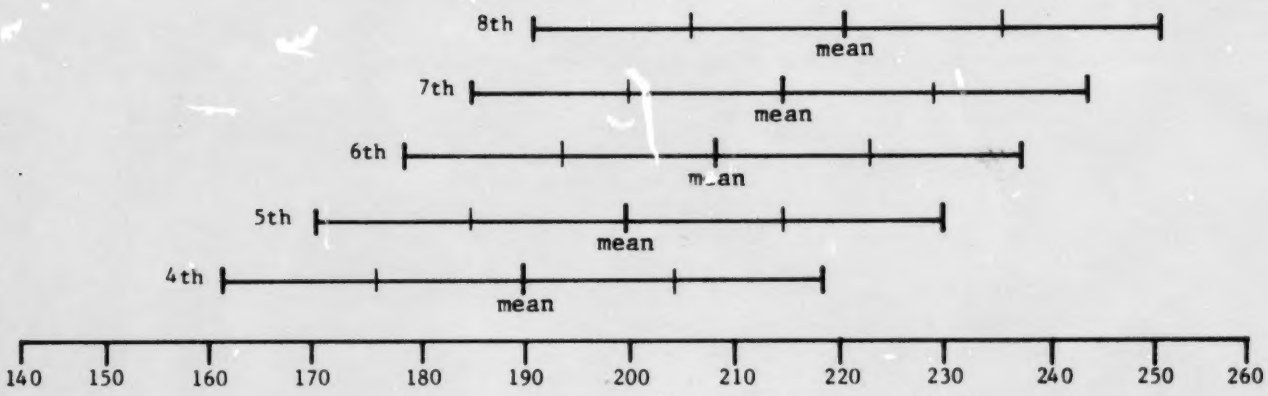
-Because all items have assigned difficulty levels, it is possible to substitute items and objectives without changing the comparability of test scores.

For instance, students may be given a pre-test consisting of one set of items of a particular difficulty level and a post-test consisting of a different set of items with a similar difficulty level. The scores on the two will be comparable because the difficulty of the items is similar.

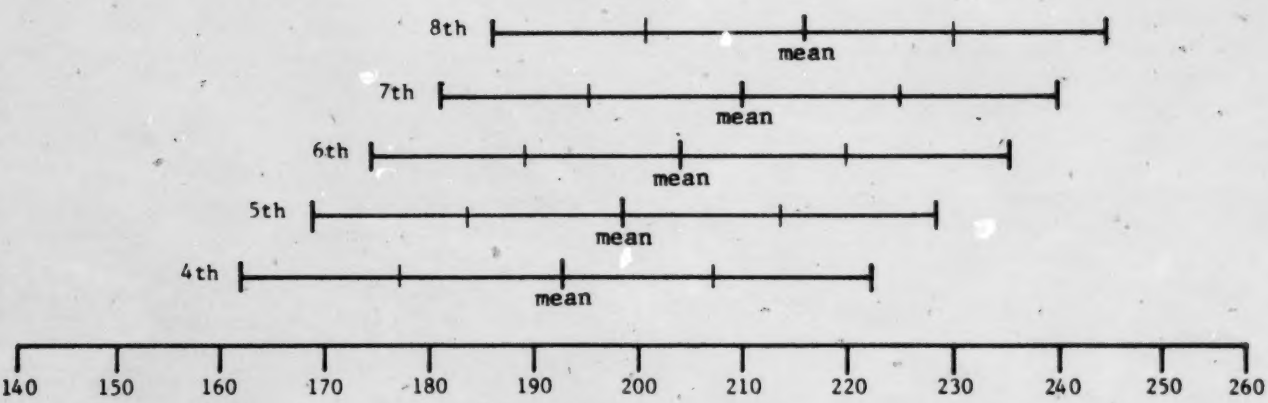
*For a detailed discussion of the Rasch method, see Benjamin D. Wright, "Solving Measurement Problems with the Rasch Model", Journal of Educational Measurement, Volume 14, No. 2, Summer 1977, pp. 97-115.

Figure 1
Average RIT Scores for Reading and Mathematics*

R E A D I N G



M A T H E M A T I C S



*These are averages for students in the Portland, Oregon school system. The average scores of Wisconsin students have not been determined, but should be nearly the same.

-The difficulty levels of all items in the item bank are calculated and placed along one continuous scale. This allows schools to construct tests which effectively measure student progress over time, not only within a single school year, but also over several years. Other testing approaches are not as effective in measuring growth over a period of time, as is shown in the following example.

If a student averages 70 percent correct on an objective-referenced reading test taken in the fall and 75 percent correct on another objective-referenced reading test taken the following spring, it appears as though the student has gained some knowledge because the test score improved. However, this assumption may be inaccurate because the relative difficulty levels of the tests are unknown. For example, if the spring test was easier than the fall test, the student may have gained little or no knowledge between tests, even though the scores show improvement. The student may simply have gotten a higher score because the test was easier.

In contrast, tests developed from the item bank allow schools to more accurately measure student progress within a school year and over a period of several years. This is possible because a student's RIT score is based not only upon the number of right or wrong answers, but also upon the difficulty of the items.

Type of Tests

The tests that can be constructed from the item bank are objective-referenced tests which are designed to measure performance in terms of specific objectives or behaviors. These objectives are listed in a series of Course Goals books. All items are multiple choice.

Development

The item bank has been developed over a period of years, beginning in 1971. It is being continually updated and expanded. The Course Goals and the original items in the bank were developed by a consortium of school districts in Portland, Oregon. Wisconsin is now one of several users of and contributors to the item bank.

All of the items in the bank have been field-tested on 300-500 students.

Content Areas

About 1500 mathematics items and 1000 reading items are presently contained in the item bank, and approximately 1000 language arts items will be added shortly. These items are suitable for most students in grades 3-8. As the bank is further developed, items appropriate for grades 9-12 will also be included.

Objectives

The objectives incldued in the Course Goals, upon which the items are based, are too extensive to list in this manual. For information on how to obtain a copy of the Course Goals book for any of the previously mentioned content areas, interested schools may contact the Pupil Assessment Office.

Using the Item Bank

The process for the construction of a customized test from the item bank involves several steps.

STEP 1: DEVELOP OBJECTIVES

Each school that wishes to construct a test from the item bank must first develop its own objectives to be assessed.

STEP 2: MATCH OBJECTIVES TO COURSE GOALS

The objectives determined by each school must then be "translated" or matched to the objectives listed in the Course Goals books which are coded to the items in the bank. This can be done in one of two ways.

a. The school can obtain a copy of the Course Goals for the subject area(s) being tested, and match their own objectives to the objectives listed in the book. This is generally the most direct, fastest, and most efficient way of proceeding.
b. The school's list of objectives can be sent to the Pupil Assessment office to be translated into the Course Goals. Because of limited staff, this is the least efficient method of proceeding.

STEP 3: SELECT ITEMS

After the Course Goals have been identified, the next step is to select items which have been designed to test these goals. This part of the process is done by Pupil Assessment staff. A number of items, from which the school may choose, will be selected. In order to select appropriate items, it is necessary for the school to answer the following questions.

a. What is the purpose of the test? Will it be diagnostic or minimum competency?
b. What grade(s) will be tested?
c. What difficulty level is desired? This will be partially dependent upon the answers to (a) and (b).
d. How many items per objective are desired? This will depend on the total test length desired, the number of objectives being tested, and the number of available items in the bank. (In a class period of 50 minutes most students will be able to respond to about 35-45 items.

STEP 4: REVIEW ITEMS

A Pupil Assessment staff member will meet with school personnel for a review of the items selected. The school personnel will determine which items are most appropriate to use.

STEP 5: FORMAT TEST BOOKLETS

When the items have been reviewed and selected, they will be arranged into a test booklet format, the items and pages numbered, and a master copy delivered to the school. This will be done by Pupil Assessment staff.

STEP 6: PRINT COPIES OF TEST

It is the responsibility of the school to have copies of the test printed. These then become the permanent property of the school.

Scoring Reports

Schools constructing customized tests from the item bank can receive three types of scoring reports:

1. District-School Profile--lists all objectives and the items measuring each objective; indicates, for each item, the percent of students who select the correct response.

2. Individual Pupil Profile--indicates how well students perform on each objective and each item; if performance standards are set, shows whether or not the student meets the standards (see "Setting Performance Standards" on page 19).

3. Item Tally Profile by School and District--shows percent of students selecting each possible response for each item.

General Testing Information

Sampling or Testing All Students

Whether a school chooses to test all students or select a random sample is dependent on at least two factors. One consideration should be the number of students in the "population" group (for example, the fourth grade). If the number is small, the required sample size will need to be nearly the same as the "population". In this case, there would be little to gain by sampling. For example, if the number of students in a grade is 60, a sample size of 52 would be needed. Table 2 indicating the sample size requirements is shown on the next page.

A second consideration is whether a school wishes to obtain information about every student or whether a general group profile is desired. If the need is for information about the group as a whole, and the number of students is fairly large (greater than 150) a sample may be the best approach. Procedures for selecting random samples can be found in most introductory statistics books.

Setting Performance Standards

Schools and school districts using the statewide tests or constructing customized tests have the option of setting performance standards for each objective. A performance standard answers the question, "How much is enough?"

There are two performance level choices: "Acceptable" and "Mastery". If performance standards are set, the computer printouts will show how well students, as a group, have met the standards. In addition, if individual pupil reports are requested, the computer printouts will indicate how well each student has met the pre-determined performance standards.

Shown in the following pages are two basic ways to set performance standards, labeled Methods I and II. In addition, there are brief summaries of two other methods which might be used to set performance standards. These are labeled Methods III and IV. Unlike the first two methods, which indicate how to set standards for specific objectives, Methods III and IV show how to set performance standards for the entire test.

WHAT CHOICES ARE THERE IN SETTING STANDARDS?

There are four options in setting performance standards:

1. Schools may choose to set performance standards for each objective.
2. Schools may choose to set performance standards for selected objectives.
3. Schools may choose not to set performance standards for any objectives.
4. Schools may choose to set performance standards for the total test, or for selected parts of the test.

NOTE: For each of the four OPTIONS schools have the choice of establishing levels of "Acceptable" or "Mastery". Both levels need not be specified for each option. For example, a school might choose to use both levels with one group of objectives, an "Acceptable" level only with a second group, and for a third group not to set any performance standards.

Table 2
DETERMINING SAMPLE SIZE FROM A GIVEN POPULATION SIZE
N=Population Size, s=sample size

N	s	N	s	N	s	N	s
10	10	150	106	460	210	2200	327
15	14	160	113	480	214	2400	331
20	19	170	118	500	217	2600	335
25	24	180	123	550	228	2800	338
30	28	190	127	600	234	3000	341
35	32	200	133	650	242	3500	346
40	36	210	136	700	248	4000	351
45	40	220	140	750	254	4500	354
50	44	230	144	800	260	5000	357
55	48	240	148	850	265	6000	361
60	52	250	152	900	269	7000	364
65	56	260	155	950	274	8000	367
70	59	270	159	1000	278	9000	368
75	63	280	162	1100	285	10000	370
80	66	290	165	1200	291	15000	375
85	70	300	169	1300	297	20000	377
90	73	320	175	1400	302	30000	379
95	76	340	181	1500	306	40000	380
100	80	360	186	1600	310	50000	381
110	86	380	191	1700	313	75000	382
120	92	400	196	1800	317	100000	384
130	97	420	201	1900	320		
140	103	440	205	2000	322		

From Krijcie, R.V., and Morgan, D.W., "Determining Sample Size for Research Activities", Educational and Psychological Measurement, 1970, 30, pp. 607-610.

The terms, "Acceptable" and "Mastery", are indefinite in their meanings. However, "Mastery" implies that the student has learned a particular subject very well; "Acceptable" also is nebulous, but implies that the student has done a "reasonable" job. A first step a school or school district must take in setting performance standards is to reach consensus on a general meaning for each concept.

WHAT IS TO BE GAINED FROM SETTING PERFORMANCE STANDARDS PRIOR TO OBTAINING TEST RESULTS?

Some schools may choose to set performance standards after they have received the test results for their students, basing the standards on the average scores and additional information about student performance. The alternative, establishing a priori levels of "Acceptable" and "Mastery" is advantageous for at least two reasons:

1. One reason is that knowledge of test results may bias expectations. For example, if students do exceptionally well on a given objective, a school might decide to raise the levels of "Acceptable" and "Mastery". In contrast, lower scores may cause a lowering of expectations.
2. A second reason for setting performance standards prior to obtaining test results lies in the process itself. A conscientious and defensible process of setting performance standards requires a careful examination and evaluation of each test objective and test item. Awareness of student results may tempt one to be less thorough in the examination and evaluation process, and instead to rely primarily on student scores as the basis for setting performance standards.

WHAT SORTS OF GENERAL CONSIDERATIONS NEED TO BE MADE IN SETTING PERFORMANCE STANDARDS?

At the outset, it needs to be emphasized that there is no generally accepted procedure which allows one to set performance standards with ease. Each requires one to make hard decisions. What the various procedures do, however, is to provide guidelines for the setting of standards.

In addition to the above, there are other points to consider:

1. Items differ in their level of difficulty. This means that each item must be looked at in the process of setting standards.
2. Performance standards should be set at realistic levels, based on an understanding and awareness of school curricula and student experiences.
3. The process of setting performance standards is not something which is done once and then forgotten. Periodically, standards should be reviewed and adjusted.
4. Performance standards may be used to gather information on specific courses or programs. In other cases, they may be used for graduation requirements or to identify students who require remediation. If the performance standards are to be used for either of the latter two purposes, and especially graduation, it is imperative that the process be carefully planned and implemented. It is recommended that in these latter cases a large number of persons be involved.

SELECTING A COMMITTEE

Each of the methods of setting performance standards utilizes a committee of persons who are qualified to evaluate student capabilities. The selection of the committee involves two steps:

1. Reaching a decision about the size of the committee.
2. Reaching a decision about the composition of the committee.

The committee need not only consist exclusively of educators; schools may decide to include community leaders, school board members, etc. In terms of the size of the committee, the goal should be to avoid a group which is either too small or too large. If the committee is too small, it may be unrepresentative. If it is too large, it may be inefficient.

A fairly large group of perhaps twenty or more persons, which is representative of your community, may be the ideal. A large group which is diverse in its membership and viewpoints, will be more credible to both educators and non-educators in your community. However, practical considerations may preclude the selection of a large and varied group of individuals. A reasonable compromise would be 10 to 15 persons.

CHOOSING THE OBJECTIVES WHICH ARE TO HAVE STANDARDS

After reaching a decision about the size and nature of the committee, there is a second necessary step. Schools must decide which objectives are to have performance standards. One way to reach this decision is to ask the following questions of each test objective:

1. Is this objective taught, and is it one our students should know?
2. Is this objective one which is not taught directly and also one which students should not be expected to know at this time?
3. Is this objective one which is not taught directly and also one which students should not be expected to know at this time?

This step is best done by teachers or persons familiar with the school curriculum; however, non-educators could assume an important role in this part of the process.

After reviewing each objective, a school may decide to set performance standards for all. However, a reasonable approach would be <u>not</u> to set performance standards for an objective if the answer to question three is affirmative.

METHOD I

The first method of setting performance standards consists of six steps:

1. Select a committee.
2. Determine which objectives are to have performance standards.
3. Have members of the committee assign each item (question) measuring an objective a numerical value. This value indicates how difficult they feel the item is. IN THIS STEP COMMITTEE MEMBERS MAY BE ASSIGNED DIFFERENT TASKS; IT IS NOT NECESSARY THAT THEY ALL DO THE SAME THING. For example, members may review different items and objectives.
4. Have members of the committee discuss their ratings with the others. On the basis of the discussion individuals may choose to change their ratings of items. A dialogue among committee members is an important part of the process.
5. Tally the ratings of the committee members and determine the average Level of Difficulty for all items measuring each objective.
6. Set the performance standard(s).

After a decision has been reached about the size and composition of the committee and the objectives which are to have performance standards, the next step is to have each committee member assign a numerical value to the test items, indicating the Level of Difficulty. A scale of recommended values is given below:

The Level of Difficulty

Value	Description
1	A very difficult item; perhaps less than 25% of our students should be expected to answer this item correctly.
2	A moderately difficult item; perhaps 25%-50% of our students should be expected to answer this item correctly.
3	A moderately easy item; perhaps 50%-75% of our students should be expected to answer this item correctly.
4	A very easy item; perhaps 75% or more of our students should be expected to answer this item correctly.

EXAMPLE OF METHOD I

An example of how Method I might be used follows.
Items 35, 36, 37, and 38 are measures of Objective 15. The Level of Difficulty, as judged by one of the committee members, is shown to the right of each item.

EXAMPLE:

Objective Number 15

Item Numbers	Level of Difficulty
35	4
36	1
37	2
38	2

Total = 9

Average for Objective = 2.25

In this instance, item 35 has been judged as "very easy," item 36 as "very difficult" and items 37 and 38 as "moderately difficult". The average of the four items is 2.25 (9 divided by 4).

Suppose that the committee consists of ten members and that their average ratings for Objective 15 are as follows:

Committee Member	Average Level of Difficulty (Objective 15)
Individual 1	2.25
Individual 2	2.50
Individual 3	2.00
Individual 4	2.25
Individual 5	1.75
Individual 6	2.25
Individual 7	2.00
Individual 8	1.75
Individual 9	2.25
Individual 10	2.00

Total = 21.00

Average for Committee = 2.10

Based on the average ratings of the ten committee members, Objective 15 has a Level of Difficulty of 2.10. The table below shows recommended Levels for "Acceptable" and "Mastery", based on the average Level of Difficulty as calculated in this example.

The value, 2.10, is sufficiently close to the value, "2", so that reasonable performance levels for "Acceptable" and "Mastery" would be 2 and 3, respectively.

	Average Level of Difficulty	Recommended Performance Levels: Acceptable	Mastery
Very Difficult	1	1	2
Moderately Difficult	2	2	3
Moderately Easy	3	3	4
Very Easy	4	4	*

*None should be set if the "Acceptable" level is 4 items.

Problems occur when the average Level of Difficulty approximates the values, 1.5, 2.5, or 3.5. For example, if the average Level of Difficulty is 1.55, the committee members must decide whether to set the performance levels for "Acceptable" and "Mastery" at 1 and 2, respectively, or possible 2 and 3.

METHOD II

The second method of setting performance standards is a slight modification of one offered by Nedelsky in 1954*. Like Method I, Method II also consists of a six step process designed to provide guidelines for setting performance standards. The six steps are as follows:

1. Select a committee.
2. Determine which objectives are to have performance standards.
3. Have members of the committee assign each item (question) measuring an objective a numerical value. This value is based on the number of answers that an average student should be able to recognize as WRONG. IN THIS STEP COMMITTEE MEMBERS MAY BE ASSIGNED DIFFERENT TASKS: IT IS NOT NECESSARY THAT THEY ALL DO THE SAME THING. For example, members may review different items and objectives.
4. Have members of the committee discuss their ratings with the others. On the basis of the discussion individuals may choose to change their ratings of items. A dialogue among committee members is an important part of the process.
5. Tally the ratings of the committee members to determine an average Probability Value for the items measuring an objective.
6. Set the performance standard(s).

EXAMPLE OF METHOD II

In the following example, there are four items measuring an objective: numbers 9, 10, 11, and 12. One of the committee members has reviewed the items and indicated the number of answers for each item that he thinks an average student should be able to eliminate as being obviously wrong.

In the case of item 9, the number "1" has been circled, indicating that it is the opinion of this committee member that an average student should be knowledgeable enough to eliminate at least one answer out of four as being possible. This means that even if the student is forced to guess the correct answer to the question, there is a one in three chance of guessing correctly. For this reason, the figure .33, is entered in the table. This value is referred to as the Probability Value.

Item Number	Number of Answers an Average Student Should Recognize as Wrong				Probability of Randomly Guessing the Correct Answer
9	0	1	2	3	.33
10	0	1	2	3	.25
11	0	1	2	3	.50
12	0	1	2	3	1.00
					Individual's Total Rating = 2.08

Item 10 is thought to be very difficult. In the opinion of the committee member an average student would be unable to eliminate any of the answers as being obviously wrong. As a result, the student has only a one in four chance of guessing the correct answer. A one in four chance provides a Probability Value of .25.

---

*Nedelsky, L. "Absolute Grading Standards for Objective Tests", Educational and Psychological Measurement, 1954, 12, pp. 3-19.

For item 11 the value is .50, while item 12 has a value of 1.00. Item 12 is felt to be very easy, and the number "3" is circled. Because three of the answers should be recognized as wrong or "absurd", the student is left with only one reasonable choice. Thus, the Probability Value is 1.00.

Suppose that ten committee members are used to judge the four items, with the following results:

Committee Member	Individual's Total Ratings
Individual 1	2.08
Individual 2	2.00
Individual 3	2.00
Individual 4	1.99
Individual 5	1.83
Individual 6	1.83
Individual 7	2.00
Individual 8	1.83
Individual 9	1.96
Individual 10	2.00

Total = 19.52
Average for Objective = 1.95

For the ten committee members the average Probability Rating for the objective is 1.95 which can be rounded to 2.00. The table below offers recommended performance levels, based on the average of the committee members for an objective measured by four items.

Recommended Performance Levels

Average Probability Rating		Acceptable	Mastery
Very Difficult	1	1	2
Moderately Difficult	2	2	3
Moderately Easy	3	3	4
Very Easy	4	4	*

*None should be set if the "Acceptable" level is 4 items.

Items measuring an objective having a Probability Rating close to "1" are very difficult items, and the "Acceptable" and "Mastery" levels should be quite low. The recommended values are 1 and 2, respectively. For the example given above, the average Probability Rating of the committee is nearly 2. Thus, it would be reasonable to set the "acceptable" level at 2 and the "Mastery" level at 3.

As with Method I, the most difficult decisions arise when the ratings approximate 1.5, 2.5, or 3.5. Members of the committee need to decide whether to establish higher or lower levels of "Acceptable" and "Mastery".

TWO STANDARD SETTING METHODS WHICH ARE NOT REPORTABLE BY THE STATEWIDE TEST COMPUTER SCORING SYSTEM

There are two additional methods which can be used to set performance standards. These methods differ from those presented earlier in this testing services manual, in that they represent procedures to set standards for the test in its entirety, rather than for each objective on the test. The procedures described earlier (Methods I and II) are designed to set performance levels for specific objectives, which are to be returned with the answer sheets. The computer scoring system is programmed to report how well students have met predetermined performance levels.

If either Method III or IV is selected it is necessary for schools to hand-record performance levels on the computer printouts. THE COMPUTER PROGRAM PACKAGE IS NOT DESIGNED TO REPORT INFORMATION ON PERFORMANCE LEVELS FOR THE ENTIRE TEST.

METHOD III

Method III is similar to Method I, in which committee members are asked to assign a value to test items which reflects how difficult they feel the item is. The recommended scale of values is shown on page 23. In the following example, a test consists of 24 items and each item has received a Level of Difficulty score.

Item Number	Level of Difficulty	Item Number	Level of Difficulty	Item Number	Level of Difficulty
1	1	9	4	17	3
2	3	10	2	18	4
3	4	11	3	19	2
4	2	12	3	20	1
5	3	13	1	21	4
6	2	14	2	22	2
7	3	15	3	23	2
8	2	16	2	24	3

Total = 61

Average Level of Difficulty = 2.54 (61 ÷ 24)

During the process of judging the difficulty level of test items, committee members may discuss their decisions and change them. After this step is completed, there are five additional steps needed to develop an estimate for a performance standard(s).

Step 1 The difficulty values of each item are totaled. In this example, the total is 61.

Step 2 The total is divided by the number of items on the test to obtain an average Level of Difficulty. In this case, the difficulty value is 2.54: 61 ÷ 24 items = 2.54.

Step 3 The average Level of Difficulty is divided by 4.0 (the top value on the four point Level of Difficulty scale). The result is a percentage:

$$\frac{2.54}{4.00} = 64\%$$

Step 4 The number of items on the test is then multiplied by the percentage obtained in Step 3. The resulting "Acceptable" performance score is 15.36: .64 X 24 items = 15.36

Step 5 The "Acceptable" performance scores, determined for each committee member, are then averaged. In the following example, the average is 15.71. This value represents the number of items the committee expects the average student to answer correctly.

Committee Member	"Acceptable" Performance Score
Individual 1	15.36
Individual 2	17.65
Individual 3	16.90
Individual 4	15.66
Individual 5	14.50
Individual 6	15.00
Individual 7	15.48
Individual 8	17.00
Individual 9	14.56
Individual 10	15.00
	Total = 157.11
	Average for the Committee = 15.71

The value, 15.71, should be rounded to 16 items. In effect, this value represents an average Level of Difficulty for the test as a whole and would represent a reasonable choice for an "Acceptable" performance score. Should a school wish to set a "Mastery" level, committee members would need to select a higher score. In this example, the score determined would be somewhere between 17 and 24.

METHOD IV

Method IV is similar to Method II. Each member of the committee is asked to judge the probability of an "average" student randomly guessing the correct answer. For each item a committee member is asked to indicate the number of answers that an average student should be able to eliminate as possible answers. The calculated value represents the probability that a student would randomly guess the correct answer.

An example is given for a test consisting of 32 items.

Item Number	Probability Value	Item Number	Probability Value	Item Number	Probability Value
1	.25	12	.25	23	.50
2	.33	13	.25	24	1.00
3	.50	14	.50	25	.33
4	.33	15	.33	26	.25
5	.50	16	.50	27	.25
6	1.00	17	1.00	28	.50
7	.25	18	.25	29	.33
8	.33	19	.33	30	.33
9	.50	20	.25	31	.50
10	.33	21	.50	32	.25
11	.50	22	.33		

Individual's Total Rating = 13.55

It is obvious that in the opinion of this committee member the test is very difficult. After an average student has eliminated the "obviously wrong" answer choices, he or she would have a high probability of guessing the correct answers to only 13 or 14 items out of the 32.

The next step is to determine the average Total Rating of all committee members. In the example below, the average is 13.28.

Committee Member	Total of Probability Values
Individual 1	13.55
Individual 2	12.90
Individual 3	14.15
Individual 4	12.05
Individual 5	13.00
Individual 6	14.50
Individual 7	12.75
Individual 8	12.80
Individual 9	13.70
Individual 10	13.40

Total = 132.80

Average for Committee = 13.28

The average rating determined by the ten committee members should be rounded to 13 items. This value could serve as the "Acceptable" level for the 32 item test. Determination of a "Mastery" level could be set as a result of a discussion among committee members. An alternative would be to substitute the word, "exceptional" for "average" student and repeat the process outlined above.

A Concluding Remark Concerning Performance Standards

The methods outined in this manual represent examples of ways to establish performance standards. Schools may, of course, modify them or develop their own procedure. Regardless of the direction taken, it is critical that the process of setting performance standards be systematic and thorough. Generally, the process itself will have more credibility if larger numbers of persons are involved. Though all the examples in this paper consist of ten persons, this committee size represents a minimum.

Utilization and Dissemination of the Test Data

In developing effective testing programs, schools should consider how they might make use of test data. The following are several questions which should be considered.

SOME GENERAL CONSIDERATIONS

1. Who is to do the basic analysis of the test results?

   -is this to be done by a committee?
   -if so, what is to be the composition of the committee?
   -only teachers and administrators?
   -curriculum specialists?
   -if a committee is formed, when is it to meet? where? for what length of time?

2. What groups will receive the analysis?

   Possible target audiences include:

   -school board members
   -mathematics, reading, economics, business education, home economics, and language teachers
   -other staff members
   -members of the administration
   -students who took the test
   -parents of students who were tested
   -the general public
   -others

3. What is to be the content of the report?

   The nature of the report is dependent upon the target audience(s) in terms of the following:

-what is to be included in the report?
-how detailed is the report to be?
-how sophisticated is the report to be?
-when is the report to be given, released?
-how is the report to be given?
    -a written report? an oral report? a press release?
    -a special media presentation (slides, for example)?
    -others, including combinations of the above?
-is the report to include recommendations?

THREE POSSIBLE TARGET GROUPS

As indicated above, there are a number of possible groups that might receive the analysis. Considerations in reporting to three of these groups are listed below.

1. Teachers

Teachers would be most interested in a report focusing on specific items. Item analysis would direct attention to the following questions:

-which items were most difficult for the students?
    -does there appear to be a pattern?
    -do the missed items give any indication of a program weakness?
    -is the objective being taught?
        -if so, at what level?
        -is there periodic and frequent reinforcement?
    -is a weakness area a priority area?
        -should it be?
    -is the weakness only at the school level, or is it also a problem area at the district and state levels?
    -is the weakness a result of misunderstanding the entire concept, or does it appear to be a fragmentary error (for example, incorrect decimal placement)?
-which students did best? which did worst?
    -do the scores on the test correlate positively with grades?
        -if not, why?
-which students missed the test?
    -are those students who missed the test different in any way from those who were present (for example, are they predominately students who do poorly in the subjects tested?)
-which items (areas) show strength?
    -are the averages for the school or district better than the state averages?
-were there any special circumstances during the day of testing that might have influenced the test results?
    -were students motivated to do their best?
    -were the students notified beforehand about the testing?
    -were the parents notified?

2. School Board Members

A report to members of the school board probably should be more general and not focus on specific test items. Important questions include:

-who should make the report?
-when should it be given?
-what should be the nature of the report?
  -should it be oral, written, or both?
  -should examples of items be included?
  -should the report emphasize comparisons?
-if the report includes a section on comparison, what are some important considerations that should be made?
  -what is to be gained? are there positive outcomes?
  -are there risks?
    -will specific schools (and teachers and classrooms) be identified?
    -will comparisons be made of students having different backgrounds (different socioeconomic backgrounds, for example)?
    -will comparisons be made of students having different abilities in the subjects tested?
    -is there a danger of "labeling" students or stereotyping?
    -should comparisons include a consideration of the problems associated with measurement, including test validity, test reliability, standard errors?
-should the report include recommendations?

3. Students and Their Parents

-should the results be placed in the student's file?
-should students/parents be given a formal presentation? a written copy of any report?
-should the test results be gone over with the student with a copy of the test at hand?
-should each student and his/her parents participate in a conference with a teacher and/or guidance counselor?
-should parents/students be made aware of some of the problems associated with measurement?
-on the basis of the test, should recommendations be made to the students/parents?